Transfer matrix method for enumeration and generation of compact self-avoiding walks. I. Square lattices

A. Kloczkowski and R. L. Jernigan
Molecular Structure Section, Laboratory of Experimental and Computational Biology, NCI, NIH, Building 12B, Room B116, MSC 5677, Bethesda, Maryland 20892-5677

(Received 4 February 1998; accepted 27 May 1998)

The transfer matrix method has been developed to enumerate and generate compact self-avoiding walks in two dimensions on the square lattice within rectangular strips of size $m \times n$. The method is significantly superior to the traditional method of computer generation of self-avoiding walks, because it is attrition-free, i.e., each computation leads to successful conformations, with no failures. The method is generalized to irregular shapes, and the extension of the method to the Monte Carlo sampling of the compact conformational space is proposed. Application of this new method to protein conformation generation is discussed, with the possibility of including several types of constraints. © *1998 American Institute of Physics.* [S0021-9606(98)50533-8]

INTRODUCTION

The generation and enumeration of all possible conformations of macromolecules is one of the most important and fundamental problems in polymer science. Various approaches have been taken. The most common requirement has been for a representative set of conformations of a random coil, which usually are taken to be on a lattice of one type or another. The simplest representation of a random coil polymer is a random walk on a lattice. In this case the excluded volume of the chain is neglected (the chain is phantom) and the number of possible walks N_{n_0} of n_0 steps on the lattice with the connectivity number z (the number of nearest neighbors to each site) is

$$N_{n_0} = z^{n_0}. \tag{1}$$

The distribution of the end-to-end distance for extremely long ($n_0 \to \infty$) phantom chains is Gaussian.

The simple random walk, neglecting excluded volume is however not a realistic model for the polymer chain. A much better approximation is the random self-avoiding walk on a lattice, where any lattice site cannot be visited more than once. The condition of self-avoidance complicates the mathematical treatment of the problem, and there is no simple analytical formula [similar to Eq. (1)] relating the number of possible conformations (walks) N_{n_0} to the length of the chain n_0. Actually, there is some evidence that the self-avoiding walk problem belongs to the ''unsolvable'' class of mathematical problems, i.e., there is no solution in terms of *D*-finite functions (usual functions of mathematical physics).[1]

For shorter chains the usual practice has been to enumerate completely self-avoiding walks. For example enumerations of self-avoiding walks on the square lattice up to 51 steps have been published.[2] The number of conformations grows exponentially with the length of the chain, and the exact enumerations for longer chains are computationally limited. It has been found that for long chains ($n_0 \to \infty$) the number of possible conformations scales with n_0 as

$$N_{n_0} \approx n_0^{\gamma-1} \mu^{n_0}, \tag{2}$$

where γ is an universal exponent dependent only on the dimensionality of the system, while μ is a lattice dependent constant which satisfies the inequality $z/2 < \mu < z-1$. Renormalization group theory calculations predict (nonrigorously) the exact value of the exponent γ to be $43/32 = 1.343\,75$ in 2D and estimate the value of γ to be 1.1615 ± 0.0011 in 3D.[3]

For longer chains, where the complete enumerations of conformations are impossible, the common practice in chain generations has been to utilize various Monte Carlo approaches to sample the conformational space.[4] In addition another approach was used in which generator matrices were utilized to calculate average properties of all conformations of a random coil.[5]

Globular proteins differ from random coils since they have dense, compact cores as a result of the substantial segregation between hydrophobic and polar residues. Because of their dense cores, compact self-avoiding walks (chains) on lattices provide an excellent model for globular proteins.[6–18] A compact self-avoiding walk is defined here as a self-avoiding walk within a compact shape, such that all sites within the shape are occupied; there are to be no voids.

The compact self-avoiding walks are essentially equivalent to the mathematical problem of Hamiltonian paths (Hamiltonian walks). A Hamiltonian walk over a graph visits all points of graph exactly once. There are no sites left unvisited, and each site is visited only once. A Hamiltonian walk that returns to the starting point is called Hamiltonian circuit. Figure 1(a) shows an example of a Hamiltonian path, and Fig. 1(b) a Hamiltonian circuit on the square lattice, both within the 4×7 rectangle.

It has been shown that such a dense compact state has large reductions in the number of conformations. The computer enumerations of compact self-avoiding walks have been reported by various authors.[7,9,10,14,15,19–21] For long chains there is a relation between the number of compact conformations $N_{n_0,p}$ and the total length of the chain n_0,

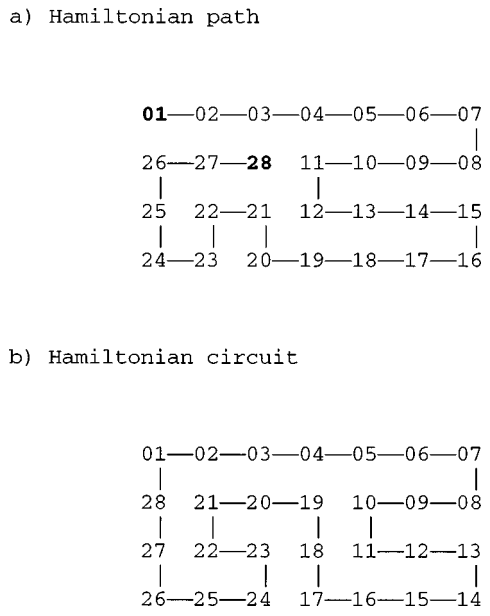

FIG. 1. Examples of a Hamiltonian path (a) and a Hamiltonian circuit (b) within the 4×7 square lattice.

with p sites located on the exterior (boundaries) of the compact shape

$$N_{n_0,p}\approx\mu_0^p\kappa^{n_0}, \tag{3}$$

where $\kappa\geqslant1$ is a connective constant, and $\mu_0\leqslant1$ is a constant [different from μ in Eq. (2)] accounting for the effect of the periphery of the shape. It has been shown that in the limiting case ($n_0\rightarrow\infty$) when Hamiltonian circuits homogeneously cover the whole lattice the connective constant κ may be approximated by

$$\kappa=z/e, \tag{4}$$

where z is lattice connectivity number ($z=4$ for the square lattice) and $e=2.71828\ldots$ is the Euler constant.[3]

The generation and enumeration of all, or of as many as possible, compact conformations is a critically important problem for protein folding. The native conformations of proteins are compact and unique. The essence of the protein folding problem is to find, for a given sequence of amino acids the most favorable conformation. This search for a unique form means that random search methods will often fail; complete enumerations, whenever feasible, are preferable. Popularly, this has been termed looking for a needle in a haystack. The number of possible conformations even for small proteins containing, say 50 residues, is extremely large. Even if we use a lattice and restrict the conformational space by searching only for those conformations within a specified volume and shape the problem is still a computational challenge for longer chains.

The standard method for generating compact conformations on lattices are self-avoiding walk calculations, either with enumerations or Monte Carlo methods. With the self-avoiding walk method it is theoretically possible to generate and enumerate all possible compact conformations within a given volume, but because of the fast growth in the number of possible conformations and because of the high attrition due to the self-avoiding (excluded volume) restriction, the time required for computations grows geometrically with the length of the chain, and becomes prohibitively long for chains of more than 30–50 bonds, the difficulty depending on the specific details of the lattice used.

The most important obstacle to the computer generation of random coil conformations is attrition. We may define the attrition as the ratio of the number of accepted (satisfying self-avoidance condition) random walks N_{sa} to the total number of generated random walks N_0. It has been shown that this ratio decays exponentially with the length of the chain n_0

$$N_{\mathrm{sa}}/N_0=\exp(-\lambda n_0), \tag{5}$$

where the attrition constant λ depends on the type of lattice. Generally λ is larger, and the decay is faster, for lattices with a smaller coordination number z, and for lower dimensions (2D rather than 3D). Hemmer and Hemmer showed that an average random self-avoiding walk on a square lattice lasts 71 steps, and will terminate due to self-avoidance.[22]

For compact structures, such as proteins the attrition problem is even more severe and limits all studies of protein folding. For example in a computer generation of compact self-avoiding walks (Hamiltonian paths) on rectangles of size $6\times n$ on the square lattice we found[21] that for $n=2$ the percentage of steps leading to successful self-avoiding conformations is 12.5%, but for $n=5$ only 0.99%, and for $n=7$ only 0.17%. (The computer program was terminating the count of the number of steps of the walk when a dead-end due to the self-avoidance was encountered.) We may estimate that for $n=14$ only about 1 step in 10^6 steps leads to a successful conformation, and for $n=20$ only 1 step in approximately 10^9 steps leads to a self-avoiding conformation. Of course for lattices with a higher coordination number z, like the cubic lattice ($z=6$), the attrition is smaller than for the square lattice ($z=4$), but it still is a major obstacle to the efficient generation of compact conformations.

For Monte Carlo methods a random generation of conformations is utilized, but the use of this method for highly compact states is quite limited, principally because of the high attrition. In addition, since the native conformation is unique, there remains a finite chance of never locating it with such a random sampling of conformational space.

All previous approaches to generating conformations for this state have encountered inefficiencies became of dead ends and unoccupied, isolated holes.[14] The present approach overcomes all of these difficulties in a direct way.

THE TRANSFER MATRIX METHOD

A new approach to the generation and enumeration of compact conformations on lattices that avoids the high attrition problem is the transfer matrix method. This method was originally proposed in the early 1980s by Derrida[23] and Klein,[24] and used by Schmalz *et al.*[25] for enumerations of Hamitonian circuits within rectangles on the square lattice.

We will first describe briefly this method and demonstrate its application to the generation and enumeration of Hamiltonian circuits on the square lattice in 2D. Then we present the extension of this method to Hamiltonian paths

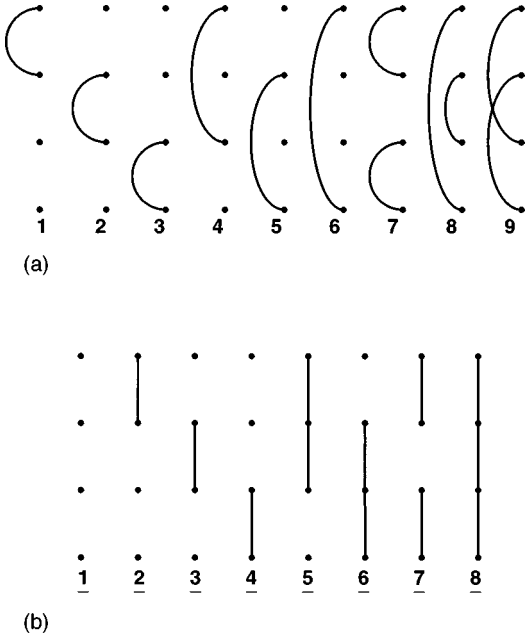

FIG. 2. The connectivity states (a) and bond distributions (b) for the generation of Hamiltonian circuits within $4\times n$ rectangles on the square lattice.

(chains) with two ends on the square lattice. We will also discuss the extension of this method to irregular compact shapes and the possibility of the sampling of the conformational space with this approach. In a subsequent paper we will develop the transfer matrix method for the generation and enumeration of both Hamiltonian circuits and Hamiltonian paths in three dimensions on the cubic lattice. Some of these important results were recently reported in a Brief Communication.[26]

The main simplifying idea in the transfer matrix method is to take individually each column of sites inside the rectangle on the square lattice and define as ''states'' the connectivities of these sites through the part of the chain on one side, here the left side of the given column. (The left side is a convention that we use, by assuming that the rectangle is being built starting from the left to the right. Also the choice of columns instead of rows is a convention.) The connectivity state is then defined as a set of graphs joining all pairs of sites in a given column connected through the chain externally to the left. With such a definition of ''state'' there are relatively few allowed ''transitions'' from a given state to the neighboring states, and these are easily specified in a general way as will be shown below.

To illustrate this method let us consider, as an example the enumeration of Hamiltonian circuits on a square lattice constrained to the $m\times n$ rectangular strip having width $m=4$ and variable length n. Figure 2(a) shows all possible external connectivities to one side (left) of the 4 sites in a column. Figure 2(b) shows all possible distributions of bonds among the 4 points on a line, including the case having no bonds at all (#1). We note that intersecting connectivities such as number 9 in Fig. 2(a) are not allowed in two dimensions. For 4 points there are thus 8 connectivity states (nonintersecting connectivities). However two of these states [numbers 4, and 5 in Fig. 2(a)] are not possible for circuits on $4\times m$ rectangles, for parity reasons (which is explained later), so the actual number of possible connectivity states is further reduced to 6. The condition that connectivities cannot intersect will lower significantly the number of allowed states for larger widths (cross sections) of the rectangle. The number of possible distributions of bonds between m sites in a column is 2^{m-1}.

To better understand the idea of connectivity states and and bond distributions let us consider the Hamiltonian circuit on the the 4×7 lattice shown in Fig. 1(b). The vertical bond distributions (starting from the left) are 8, 3, 4, 6, 3, 1, 7 [from Fig. 2(b)]. The connectivity states [from Fig. 2(a)] starting from the left are 6, 8, 1, 6, 8, 8 respectively. The first connectivity state (starting from the left) is state number 6 [from Fig. 2(a)] because in the second column of sites only sites 02 and 25 are connected through the chain on the left. The next connectivity state is state number 8 [from Fig. 2(a)] because this corresponds to the left-side connectivity of sites 03 and 24, and 20 and 23 in the third column of sites, etc.

Before illustrating the method of the construction of the transfer matix we discuss the problem of counting the number of connectivity states for Hamiltonian circuits.

COUNTING CONNECTIVITY STATES FOR CIRCUITS

The main limitation to the transfer matrix method is caused by the rapidly growing number of connectivity states as we increase the width (cross section) m of the rectangle.

For $m=4$ there are 8 nonintersecting connectivity states as seen in Fig. 2(a), but this number was further reduced to 6 by taking account of parity considerations.

The important problem is to evaluate the number of possible connectivity states for compact circuits on the square lattice for any width m of the rectangular strip. The number of all possible different pairwise connectivities of m points on the line, where m is an even number, and *all* points are pairwise connected but each point can be connected with only one other point, and additionally intersections of connectivities are allowed is given by the formula

$$M_m=(m-1)(m-3)\ldots 1=(m-1)!!. \tag{6}$$

For example for $m=6$ there are $5!!=5\cdot 3\cdot 1=15$ different possibilities as shown in Fig. 3(a), and for $m=4$ there are $3!!=3$ [connectivity states 7,8,9 in Fig. 2(a)]. For odd numbers of points m the number of possible pairwise connectivities of $m-1$ points equals $(m-2)!!$ and there are m possibilities of chosing one unconnected point, so $M_m=m!!$. This is illustrated by Fig. 3(b) for $m=5$.

In two dimensions intersecting connectivities are not allowed. For example for the $m=6$ points shown in Fig. 3(a) out of the 15 possibilities only 5 connectivities fulfill this condition [numbers 1, 3, 7, 13, and 15 in Fig. 3(a)]. It is possible to derive the recursive formula for the number M_m of possible connectivities of an even number of m points when the intersecting connectivities are eliminated. First we should note that for an even number of points m the connectivities between points separated by an odd number of sites are not allowed, because it would lead to intersecting connectivities in the process of completing the connectivities of all m points.

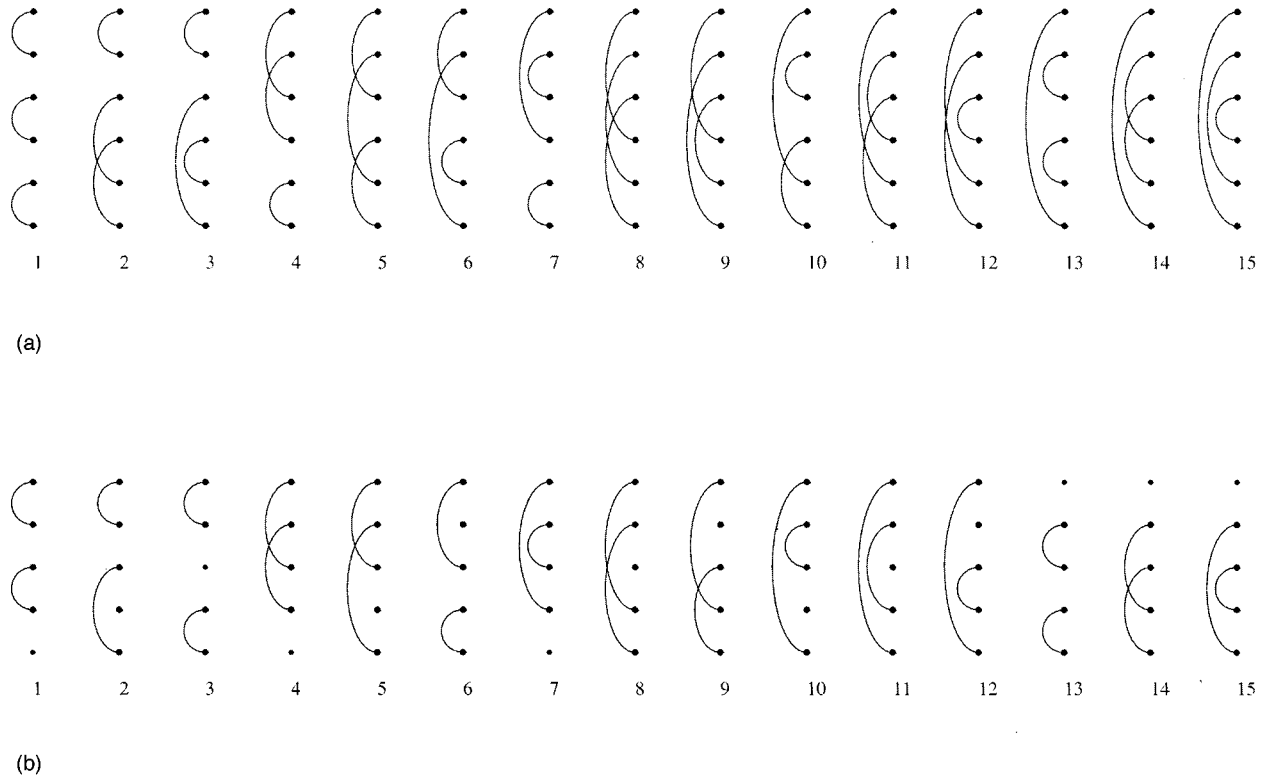

FIG. 3. All possible different pairwise connectivities of m points on the line, where maximum possible points are pairwise connected and intersections of connectivities are allowed. The cases of an even number of points $m=6$, and an odd number of points $m=5$ are shown in Figs. 3(a) and 3(b) separately.

Let us, for illustration, consider the case of $m=8$ points shown in Fig. 4. The first point can be connected with the 2nd, 4th, 6th, and 8th. The remaining points can be connected as in M_6, M_2M_4, M_4M_2, and M_6 different ways, respectively. This leads to a simple recurrence formula for the number of possible non-intersecting connectivities between $m=2\nu$ points,

$$M_{2\nu}=\sum_{k=0}^{\nu-1} M_{2\nu-2-2k}M_{2k} \tag{7}$$

with the convention $M_0=M_2=1$. Equation (7) is similar to the definition in combinatorics of the Catalan numbers C_n,

$$C_n=\sum_{i=1}^{n-1} C_iC_{n-i}=\frac{1}{n}\binom{2n-2}{n-1} \tag{8}$$

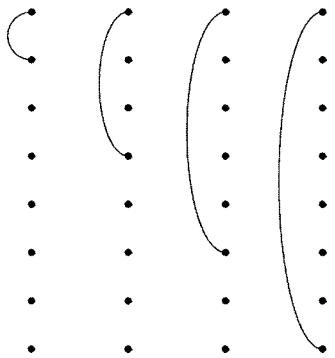

FIG. 4. The illustration for the derivation of Eq. (7). All possible connectivities of the uppermost point to other sites in the column for $m=8$ points.

with the convention $C_0=0$, $C_1=1$. The relation between $M_{2\nu}$ and the Catalan numbers leads to the compact formula for $M_{2\nu}$ shown below

$$M_{2\nu}=C_{\nu+1}=\frac{1}{\nu+1}\binom{2\nu}{\nu}. \tag{9}$$

Table I compares the numbers of different possible pairwise connectivities of all $m=2\nu$ points with intersections allowed and for all possible pairwise nonintersecting connectivities calculated from Eqs. (6) and (9), respectively. Table I shows that the condition of the elimination of intersecting connec-

TABLE I. The number of pairwise connectivities of an even number m of sites in a column, such that all points are connected and each point is connected only once. The first column shows the number of such possibilities M_m calculated from Eq. (6) when intersections are allowed, and the second column shows M_m calculated from Eq. (9) when intersecting connectivities are not allowed.

m	$M_m=(m-1)!!$ intersections allowed	M_m [Eq. (9)] intersections eliminated
2	1	1
4	3	2
6	15	5
8	105	14
10	945	42
12	10 395	132
14	135 135	429
16	2 027 025	1 430
18	34 459 425	4 862
20	654 729 075	16 796

TABLE II. The number of possible connectivity states for m points distributed along the line for varying m. Here N_m^{circ} and N_m^{chain} denotes the numbers of connectivity states for Hamiltonian circuits, and Hamiltonian chains, respectively. The numbers $N_m^{\mathrm{circ},r}$ and $N_m^{\mathrm{chain},r}$ denote the numbers of connectivity states for Hamiltonian circuits and Hamiltonian chains reduced by eliminating states with the wrong parity.

m	N_m^{circ}	$N_m^{\mathrm{circ},r}$	N_m^{chain}	$N_m^{\mathrm{chain},r}$
2	1	1	4	4
3	3	3	12	12
4	8	6	36	34
5	20	19	105	104
6	50	32	311	293
7	126	113	924	911
8	322	182	2 766	2 626
9	834	706	8 313	8 185
10	2 187	1 117	25 072	24 002
11	5 797	4 647	75 790	74 640
12	15 510	7 280	229 494	221 264
13	41 834	31 886	695 721	685 773
14	113 633	49 625	2 110 824	2 046 816
15	310 571		6 407 756	
16	853 466		19 458 562	

tivities in two dimensions is extremely effective in reducing the numbers, especially for wide strips. In three dimensions such a condition cannot be applied, and the number of possible connectivities grows extremely rapidly; hence the enumerations of compact conformations in 3D are somewhat more difficult as will be shown in a subsequent paper.

Equation (9) gives the number of all possible nonintersecting connectivities of $m=2\nu$ points, such that all the points are connected. The number of all possible nonintersecting connectivities without the condition that all points are connected follows the binomial distribution

$$N_m^{\mathrm{circ}}=\sum_{k=1}^{[m/2]}\frac{m!}{(m-2k)!(2k)!}M_{2k}=\sum_{k=1}^{[m/2]}\frac{m!}{(m-2k)!(k)!(k+1)!}. \tag{10}$$

Here $[m/2]$ denotes the integer part of the number $m/2$ and this formula is valid both for even and odd m. For example for $m=4$ there are 8 nonintersecting connectivity states shown as the first eight states in Fig. 2(a). Six of these states have a single connection $(k=1)$, and two have a pair of connections $(k=2)$. The first column in Table II shows the numbers of all possible connectivity states N_m^{circ} for circuits on the square lattice calculated from Eq. (10) for varying m.

We should note that some of these connectivity states cannot occur because for parity reasons. We may substantially reduce the number of all possible connectivity states by removing states having impossible parities.

Let us consider a column composed of m sites. To each site of the column we assign a parity value 1 or -1. We may use a convention that sites with odd numbers: 1,3,5, . . . have parity -1, and sites with even numbers 2,4,6, . . . have parity $+1$. The parity of the connectivity state is now defined as the sum of the parities of all *unconnected* sites. For example, the parity of the connectivity state #1 shown in Fig. 2(a) is 0 and the parity of the connectivity state #5 shown in Fig. 2(a) is -2.

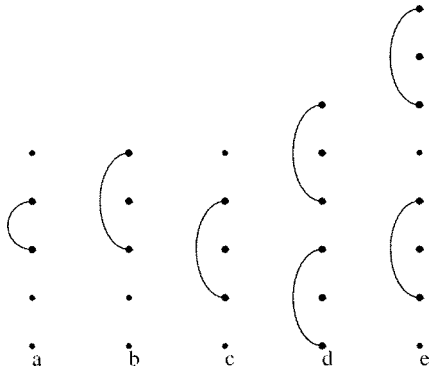

FIG. 5. Examples of physical (a), (b), (d) and unphysical (c), (e) connectivity states.

Connectivity states which having an *absolute* value of parity larger than 1 are unphysical and may be eliminated. It is relatively easy to understand this conjecture for states with single connectivity (only two sites in a column connected) and an even number m of sites in a column. The part of the rectangle for the Hamiltonian circuit corresponding to this connectivity state (the part of the rectangle to the left of the column associated with the state under consideration) always contains an even number of sites. Each point (i,j) of the rectangle has a defined parity $p(i,j)$ which is either $+1$ or -1, and which can be calculated from the formula

$$p(i,j)=2\ \mathrm{mod}(i+j,2)-1 \tag{11}$$

with $1\leqslant i\leqslant m$, and $1\leqslant j\leqslant n$, so there is a chessboardlike structure to the distribution of parities of sites within the rectangle. Each single step of the walk on the square lattice connects only sites of opposite parity, and the walk through the even number of sites in the left part of the rectangle always leads to the situation when the starting site of the walk and the ending site of the walk must have opposite parities. This means that two connected sites of the single connectivity state must also have opposite parities, i.e., the sum of parities of the unconnected sites is zero (for even m).

When we have an odd number m of sites in a column, the walk in the part of the rectangle on the left of the column of sites corresponding to the considered connectivity state can have passed through an even or an odd number of sites, depending on the number of columns. When the number of sites in that part of the rectangle is even the situation is similar to the one considered earlier, and the start and the end sites of the walk must have opposite parities [as in Fig. 5(a)]. If the number of sites in that previous part of the rectangle were odd, then the walk must start and end at sites of the same majority parity [like in Fig. 5(b)]. Figure 5(c) shows an example of an unallowed unphysical connectivity state.

The situation becomes more complicated if there is more than one pair of connected sites in the connectivity state. The connections between two sites with the same parity (e.g., $+1$) is possible if there is another connection between two sites each with parity -1, compensating the total parity of the state. An example of such a state is shown in Fig. 5(d).

Again, the absolute value of the sum of parities of the unconnected sites in the connectivity state, however cannot be larger than 1. Figure 5(e) shows an example of such an unphysical connectivity state.

The second column in Table II shows the reduced number of all possible connectivity states for Hamiltonian circuits $N_m^{\mathrm{circ},r}$ calculated for varying m, after elimination of connectivity states with wrong parities. We see that a significant number of connectivity states can be eliminated, especially for a large, even width of the rectangle m. The reduction of the number of states is much less for odd m, because states with wrong parity [such as in Fig. 5(c)] are relatively rare.

The number of all possible distributions of vertical bonds within a column of m sites for Hamiltonian circuits is much easier to calculate and it is

$$B_m = 2^{m-1}. \tag{12}$$

Figure 2(b) shows all possible bond distributions for $m=4$.

TRANSFER MATRIX FOR HAMILTONIAN CIRCUITS

To illustrate the method we consider the enumeration of Hamiltonian circuits on a square lattice for the $m \times n$ rectangular strip with fixed width $m=4$ and variable length n. The transfer matrix **T** is constructed by combining all connectivity states [Fig. 2(a)] with all bond distributions [Fig. 2(b)] for one column of sites and finding the resulting connectivity states of the next column (to the right) formed by the feasible combinations. Figures 6(a)–6(g) illustrate several simple cases. The + sign denotes the combination of connectivity states and bond distributions and the → sign indicates the product states of this combination. For example, the combination of the connectivity state number 1 [from Fig. 2(a)] and the bond distribution number $\underline{4}$ [from Fig. 2(b)] shown in Fig. 6(a) leads to connectivity state number 7 from Fig. 2(a). We can therefore define a step or ''transition'' from connectivity state 1 to state 7. Figures. 6(b) and 6(c) are other examples of allowed transitions. The combinations which lead to triple connections such as the combination shown in Fig. 6(d), to the formation of small completed loops [like the one shown in Fig. 6(e)], or leave some of the sites unoccupied [Fig. 6(f)] are not allowed. Figures. 6(g) and 6(h) show examples of combinations which complete the formation of Hamiltonian circuits.

As we stated earlier the transfer matrix **T** is constructed by combining the connectivity states with bond distributions and finding out the connectivity state of the next neighboring column of sites on the right. The element T_{ij} of the transfer matrix is zero if there is no possible transition from connectivity state i to state j. If there are possible transitions from state i to state j then T_{ij} shows the number of different ways to realize this transition. We should note that for Hamiltonian circuits on the square lattice the transitions between connectivity states are unique, so that the elements T_{ij} of the matrix **T** are either 0 or 1, but generally (for Hamiltonian circuits in three dimensions on the cubic lattice, or for chains with ends) there can be multiple ways to achieve the transition between two connectivity states, by using different distributions of bonds, and some of the elements of the transfer matrix T_{ij} can be integers larger than 1. We could of course define each state as a unique combination of a connectivity state and a distribution of bonds, so that the elements of the transfer matrix for such defined states would always be 0 or

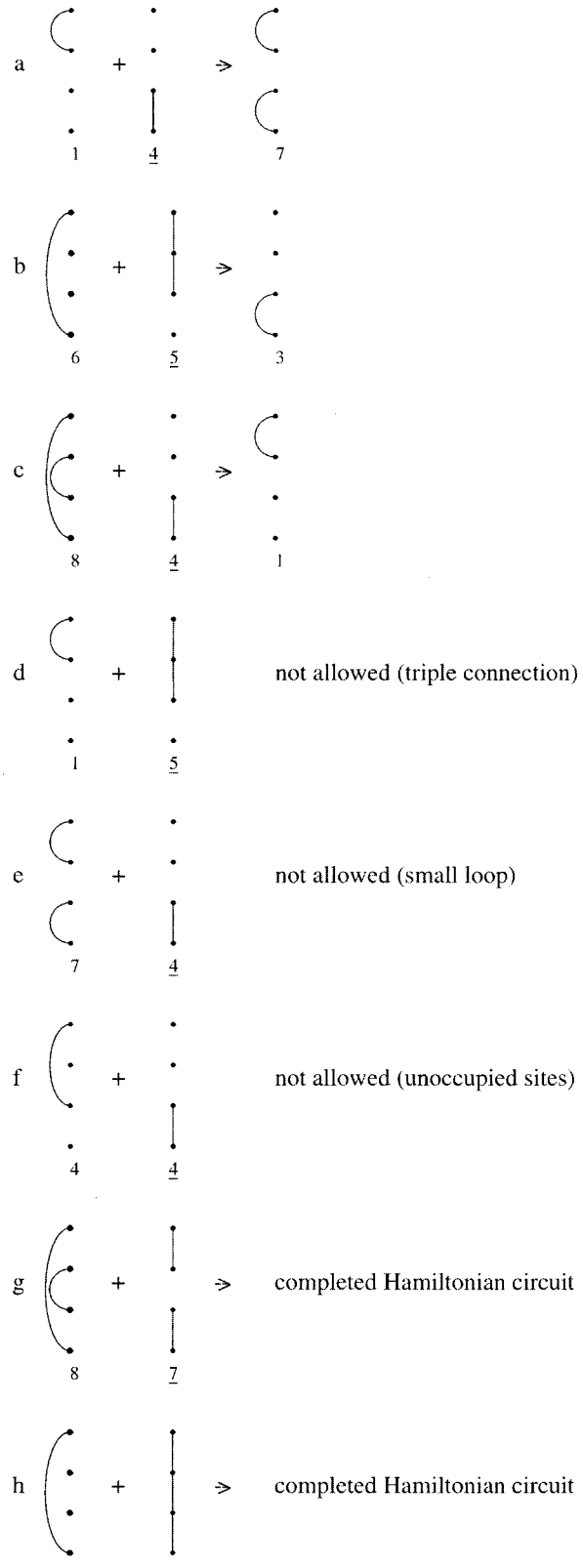

FIG. 6. Illustration of considerations entering into the construction of the transfer matrix for the enumeration of Hamiltonian circuits within $4 \times n$ rectangles on the square lattice.

1; however, such a generalized definition of states would substantially increase the size of the transfer matrix **T**.

We first construct the vector **u** of the starting states with elements u_i, for each connectivity state i [such as in Fig. 2(a)] as the first state on the left in the process of building a circuit (we use a left to right convention). The number u_i counts the number of different ways in which this may be realized. For closed circuits in two dimensions on the square lattice the elements of the vector of the starting states are either 0 or 1, but in three dimensions or for chains with ends u_i can be an integer larger than 1. As starting states we use the distributions of bonds [such as shown in Fig. 2(b)] which do not contain any unoccupied sites [#7 and #8 in Fig. 2(b)] and figure out the connectivity state to which the given distribution of vertical bonds transforms if the horizontal bonds connecting those vertical bonds to bonds to the neighboring column on the right side are added. For example distributions #7 and #8 in Fig. 2(b) transform to connectivity states 7 and 6 in Fig. 2(a), respectively.

We also construct the vector **v** of ending states with elements v_i showing how a given connectivity state i may form a closed circuit by combining it with the distribution of vertical bonds [this is illustrated by Figs. 6(g) and 6(h)]. For Hamiltonian circuits in two dimensions the components of the vector **v** are either 0 or 1, but for chains with ends (or for closed circuits in three dimensions) v_i may be an integer larger than 1. The possible ending connectivity states are states 6 and 8 in Fig. 2(a).

For example, if we remove connectivity states 4, 5 (for parity reasons) and 9 (because it is impossible in two dimensions) in Fig. 2(a) and renumber the remaining states from 1 to 6, then the transfer matrix **T**, and the vectors of the starting states **u** and the ending states **v** are

$$\mathbf{T}=\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{u}^T=(000110),$$

$$\mathbf{v}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}. \tag{13}$$

The number N_c of possible Hamiltonian circuits on the rectangle of size $m\times n$ on the square lattice is then given by the simple formula

$$N_c=\mathbf{u}^T(\mathbf{T})^{n-2}\mathbf{v}. \tag{14}$$

RESULTS FOR CIRCUITS

The enumerations of Hamiltonian circuits within rectangular strips $m\times n$ of the varying length n were first published by Schmalz, Hite, and Klein in 1984.[25] The enumerations published by these authors provided only for the cases when both n and m were even. For the case when both m and n are odd the formation of the Hamiltonian circuit is, of course, impossible and N_c equals zero. Table IV shows the results of enumerations for varying values of n and m (both odd and even) and extended to numbers N_c as large as 10^{26}. There is complete numerical agreement with the results of Schmalz, Hite, and Klein[25] and with traditional computer enumerations of Hamiltonian circuits by self-avoiding walks.[21] Figure 7 shows the plot of the number of circuits N_c vs the length n of the rectangle for varying width m of the $m\times n$ rectangle.

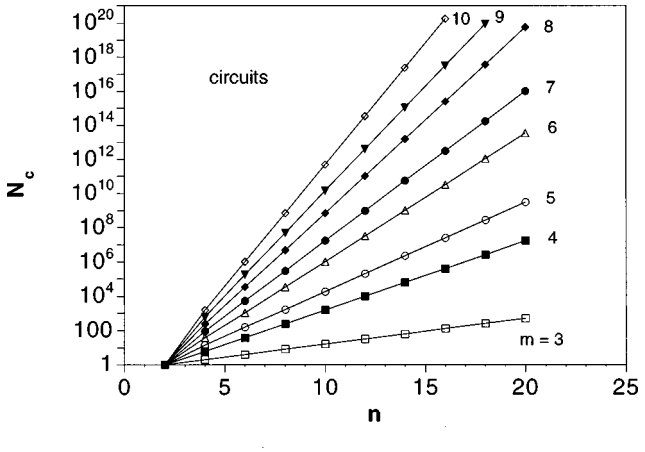

FIG. 7. The plot of the number of possible Hamiltonian circuits N_c vs the length n for varying widths m of the $m\times n$ rectangle.

The main difficulty in the extension of these enumerations to rectangular strips with larger width m, is the rapidly growing number of connectivity states as discussed earlier. The discussion about a possible approach to systems with extremely large number of connectivity states is provided in the last part of the paper.

TRANSFER MATRIX METHOD FOR HAMILTONIAN PATHS

One of the main purposes of this paper is to extend the transfer matrix method to Hamiltonian paths (i.e., Hamiltonian chains with two ends) on the square lattice. To extend the method to Hamiltonian chains in two dimensions on the square lattice we generalize the definition of the connectivity state to include all connectivities having *up to two ends*. This means that we consider individual connectivity states for circuits (no ends) presented above, and additionally we consider connectivity states with one end and two ends. Figure 8(a) shows all possible *nonintersecting* connectivities for 4 sites in a column. The connectivities to ends are represented on the connectivity diagrams as single connected left-sided lines, while the regular connectivities of two points are represented by left-sided arches joining two points. The first 8 connectivity states in Fig. 8(a) are the same as in Fig. 1(a) and contain no connections with ends. (The states 4 and 5 again do not contribute to the formation of final Hamiltonian chains and can be eliminated, due to parity reasons, as for closed circuits.) The states 9–24 contain connectivities specifying the position of one end, and the remaining states

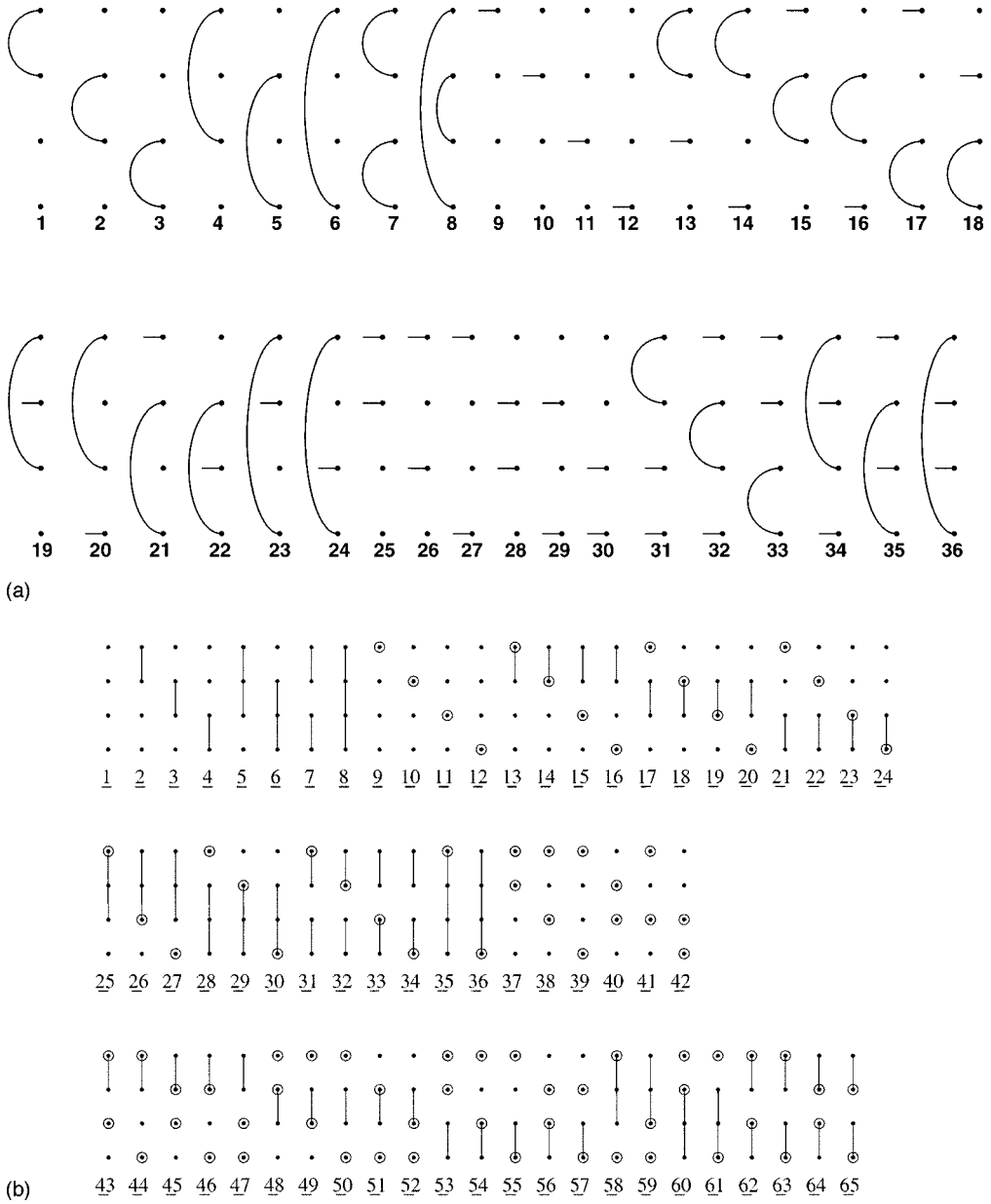

FIG. 8. The connectivity states (a) and bond distributions (b) for the generation of Hamiltonian paths (chains) within $4\times n$ rectangles on the square lattice.

25–36 contain connectivities placing two ends. Similarly we generalize the definition of the distribution of bonds by also including the end points of the chain. Figure 8(b) shows all possible distributions of *bonds and ends* of the chain. The end points in Figs. 8(b) are represented by small circles. The first 8 distributions in Fig. 8(b) are the same as in Fig. 1(b) and contain no ends of the chain. The distributions from 9 to 36 contain one end of the chain, and the remaining distributions contain two ends. We note that distributions containing the two ends attached to the same single bond (or to a sequence of connected bonds) are not allowed, because this would terminate the chain. Also the ending points cannot be placed in the middle of a sequence of connected bonds.

To illustrate better the idea of connectivity states and and bond distributions for Hamiltonian paths let us consider the Hamiltonian path on the the 4×7 lattice shown in Fig. 1(a). The vertical bond distributions (starting from the left) are 28, 4, 22, 3, 1, 1, 7 [from Fig. 8(b)]. The connectivity states [from Fig. 8(a)] starting from the left are 21, 15, 27, 32, 32, 32, respectively. The first connectivity state (starting from the left) is state number 21 [from Fig. 8(a)] because in the second column of sites site 02 is connected to an end on the left, and sites 27 and 23 are connected through a loop on the left. The next connectivity state is state number 15 [from Fig. 8(a)] because this corresponds to the left-side connectivity of site 03 to an end and left-side connectivity of sites 28 and 21 through a loop in the third column of sites, etc.

The transfer matrix **T** is constructed in the same way as for Hamiltonian circuits by combining connectivity states [Fig. 8(a)] with bonds and end distributions [Fig. 8(b)] and finding out the connectivity states of the next neighboring column on the right obtained by this combination. Figure 9 illustrates this process. For example the combination of the connectivity state 1 from Fig. 8(a) with bond distribution

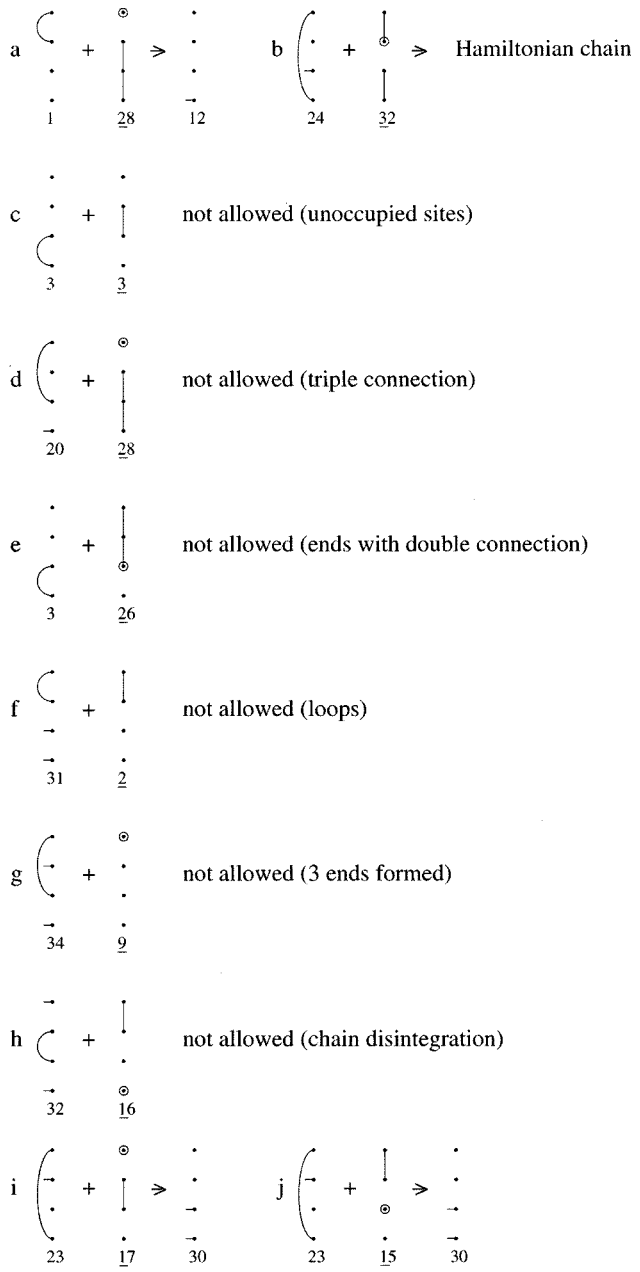

FIG. 9. Illustration of the method of construction of the transfer matrix for the enumeration of Hamiltonian paths (chains) within $4\times n$ rectangles on the square lattice.

#28 from Fig. 8(b) leads to the connectivity state 12 [Fig. 9(a)]. Similarly there are disallowed cases; triple connections [Fig. 9(d)], double connections of the ends [Fig. 9(e)], and the formation of loops [Fig. 9(f)] (including Hamiltonian circuits, since now we exclude these in the generation of chains with ends). Additionally these combinations cannot lead to the formation of more than two ends [such as in Fig. 9(g)], to the disintegration of the chain [as shown in Fig. 9(h)], or to unoccupied sites [Fig. 9(c)]. For Hamiltonian circuits on the square lattice the elements of the transfer matrix **T** are either 0 or 1, but for Hamiltonian chains with ends, integers larger than 1 are allowed. Figures 9(i)–9(j) illustrate the possibility of a given connectivity state combined with two different bond distributions that both lead to the same connectivity state. The vectors of the starting states **u** and of the end states **v** are constructed in the same way as for Hamiltonian circuits. To find out the starting connectivity states we examine all bond distributions (generalized by incorporating the end points of chains), that do not contain unoccupied sites [such as numbers 7, 8, 27, 28, 31–36, 47, 50, 53, 58–65 in Fig. 8(b)], and figure out connectivity states to which they transform if the horizontal bonds joining them with bonds in the next column on the right side are added. For example distribution 33 in Fig. 8(b) transforms to connectivity number 14 in Fig. 8(a), so state number 14 is a starting state. Some starting connectivity states are degenerate, i.e., several different bond distributions lead to the same state. For example state 27 in Fig. 8(a) is threefold degenerate, because it can be obtained from three different bond distributions, 59, 60, and 64 in Fig. 8(b). The vector **v** of the ending states is constructed by finding out which connectivity state combined with the bond distributions leads to the formation of a Hamiltonian chain. Figure 9(b) shows an example of the connectivity state leading to the formation of a Hamiltonian chain. Similarly as with starting states, ending states may also be degenerate. For example the connectivity state number 8 is doubly degenerate as an ending state because its combinations with both bond distributions 47 and 53 in Fig. 8(b) lead to the formation of Hamiltonian chains.

The problem of degeneracy could be eliminated by defining the state as a combination of the ''old'' connectivity state with one bond distribution, but this would significantly increase the size of the transfer matrix for enumerations. The enumeration of the number of possible Hamiltonian chains is carried out by employing the same equation [Eq. (14)] as for Hamiltonian circuits. Table IV shows examples of the enumeration of Hamiltonian chains on a square lattice in a rectangle of size $m\times n$ for several values of m and various n.

THE NUMBER OF POSSIBLE CONNECTIVITY STATES FOR THE HAMILTONIAN CHAINS WITH ENDS

The number of possible connectivity states for Hamiltonian circuits N_m^{circ} is given by Eq. (10). For Hamiltonian chains with ends the number of possible connectivity states N_m^{chain} contain those for Hamiltonian circuits plus additionally all different connectivity states containing one or two ends, i.e.,

$$N_m^{\mathrm{chain}}=N_m^{\mathrm{circ}}+\sum_{k=0}^{[m/2]}\left[\binom{m-2k}{1}+\binom{m-2k}{2}\right]\binom{m}{2k}M_{2k}, \tag{15}$$

where M_{2k} is given by Eq. (9) for nonintersecting connectivities.

Using the identity for binomial coefficients

$$\binom{n}{1}+\binom{n}{2}=\binom{n+1}{2} \tag{16}$$

and Eqs. (9)–(10) we obtain the following formula for the number of possible nonintersecting connectivity states for Hamiltonian chains:

$$N_m^{\text{chain}} = \binom{m+1}{2} + \sum_{k=1}^{[m/2]} \left[\binom{m+1-2k}{2} + 1 \right] \times \frac{m!}{k!(k+1)!(m-2k)!}. \quad (17)$$

The third column in Table II shows the number of all possible connectivity states for Hamiltonian chains N_m^{chain} calculated from Eq. (17) for various m. Similarly as for Hamiltonian circuits some of the connectivity states can be eliminated because of bad parity. The only connectivity states with bad parities are the same as those eliminated for Hamiltonian circuits.

The fourth column in Table II shows the reduced number of all possible connectivity states for Hamiltonian chains $N_m^{\text{chain},r}$ calculated for varying m, after elimination of connectivity states with wrong parities. We can see that the elimination of states with wrong parities for chains with ends does not lead to a significant reduction in the number of connectivity states. Such a reduction was relatively much more substantial for Hamiltonian circuits.

Another interesting problem is the comparison of the number of bond distributions for Hamiltonian paths (chains with ends) and Hamiltonian circuits (no ends). Figures. 2(b) and 8(b) show all possible bond distributions for $m=4$ points for circuits and paths, respectively. The number of bond distributions (including ends) for paths (chains) includes bond distributions with zero ends—same as for circuits, distributions with one end, and finally distributions with two ends. In the construction of bond distributions for paths ends [given by small circles in Fig. 8(b)] cannot be placed in the middle of a sequence of connected bonds. Also two ends cannot be placed at two opposite ends of one bond, or a sequence of connected bonds. The number of possible bond distributions B_m for chains depends on the width (cross section) of the rectangle m, but cannot be expressed by a simple formula such as Eq. (12) for circuits. The last column of Table III shows the number of possible bond distributions B_m (including ends) for chains for varying width of the rectangle m. The middle column in Table III shows the corresponding number of bond distributions for Hamiltonian circuits, which is given by the simple formula $B_m = 2^{m-1}$. Table III shows that the number of bond distributions for chains (with ends) increases significantly by comparison with circuits containing no ends.

TABLE III. The number of possible bond distributions for rectangles of varying width m for Hamiltonian circuits (second column) and for Hamiltonian chains (last column).

m	$B_m = 2^{m-1}$ circuits	B_m chains
2	2	7
3	4	22
4	8	65
5	16	181
6	32	482
7	64	1 240
8	128	3 104
9	256	7 600
10	512	18 272
11	1024	43 264
12	2048	101 120
13	4096	233 728
14	8 192	535 040

RESULTS FOR CHAINS WITH ENDS

Table V shows enumerations of the number of Hamiltonian chains (with two ends) N_c on the square lattice within the rectangles of size $m \times n$ for various widths of the rectangle m and variable length n. The results (N_c vs n) were plotted in Fig. 10 for varying width m of the $m \times n$ rectangle.

Similarly as in Ref. 25 we calculated connective constants κ_m, [defined by Eq. (3)] for different widths m of the rectangle. The limiting values κ_m for infinite strips for Hamiltonian chains were the same as for Hamiltonian circuits (shown in Table I in Ref. 25).

Calculations have been performed for rectangles with maximum width $m=8$. For larger m the problem becomes the increasing size of the transfer matrix. For $m=8$ there are 2766 connectivity states, and this number can be further reduced to 2626 by elimination of states with bad parity. The transfer matrix **T** is then of size 2626×2626. For example, for the case of a rectangle of size $10 \times n$ the number of the connectivity states is 24 002, and for the case of $14 \times n$ there are 2 046 816 states. Matrices **T** are sparse (most of their elements are zeros), and we could use methods developed for sparse matrices to perform calculations for rectangles wider than 8. We should note that we can also reduce the number of Hamiltonian chains by eliminating states which are related by various symmetry relations. To eliminate the symmetries associated with the shape of the $m \times n$ rectangle we divide the number of Hamiltonian chains obtained from Eq. (1) by $\sigma/2$ where σ is the symmetry number ($\sigma=4$ for rectangles with $n \neq m$ and $\sigma=8$ for squares $n=m$).

Tables IV and V show clearly how powerful is the transfer matrix method. These exact enumerations could never be obtained by using traditional methods. The exact enumerations of self-avoiding walks are important for the mathematical theory of self-avoiding walk and for the renormalization group theory. These exact results shown in Tables IV and V are also easy to verify proof that our method is free of errors.

DISCUSSION

The transfer matrix method is a powerful tool for studying protein conformations. It is extremely fast since, instead of growing conformations one bond at a time, we build the whole conformation column by column and completely avoid attrition. If each site within a column were assigned a ''letter'' of the amino acid alphabet then, assuming that we have only interactions between nonbonded nearest neighbors, all calculations of energy with the addition of a new column can be immediately performed and some blocks of columns of highest energy could be discarded.

A major advantage of the transfer matrix method is that, once the transfer matrix is calculated, enumerations can be carried out easily for any length n. The only limitation is due to specific computer limits in treating large integers. The method of calculating the transfer matrix can be easily vec-

TABLE IV. Enumerations of Hamiltonian circuits on $m \times n$ rectangles.

n	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
2	1	1	1	1	1
3	0	2	0	4	0
4	2	6	14	37	92
5	0	14	0	154	0
6	4	37	154	1 072	5 320
7	0	92	0	5 320	0
8	8	236	1 696	32 675	301 384
9	0	596	0	175 294	0
10	16	1 517	18 684	1 024 028	17 066 492
11	0	3 846	0	5 668 692	0
12	32	9 770	205 832	32 463 802	966 656 134
13	0	24 794	0	181 971 848	0
14	64	62 953	2 267 544	1 033 917 350	54 756 073 582
15	0	159 800	0	5 824 476 298	0
16	128	405 688	24 980 352	32 989 068 162	3 101 696 069 920
17	0	1 029 864	0	186 210 666 468	0
18	256	2 614 457	275 195 536	1 053 349 394 128	175 698 206 778 318
19	0	6 637 066	0	5 950 467 515 104	0
20	512	16 849 006	3 031 685 984	33 643 541 208 290	9 952 578 156 814 524

n	$m=8$	$m=9$	$m=10$
2	1	1	1
3	8	0	16
4	236	596	1 517
5	1 696	0	18 684
6	32 675	175 294	1 024 028
7	301 384	0	17 066 492
8	4 638 576	49 483 138	681 728 204
9	49 483 138	0	13 916 993 782
10	681 728 204	13 916 993 782	467 260 456 608
11	7 837 276 902	0	10 754 797 724 124
12	102 283 239 429	3 913 787 773 536	328 076 475 659 033
13	1 220 732 524 976	0	8 091 313 110 371 792
14	15 513 067 188 008	1 100 831 164 969 864	233 977 398 720 987 284
15	188 620 289 493 918	0	6 002 042 996 016 384 360
16	2 365 714 170 297 014	309 656 520 296 472 068	168 435 972 906 750 526 954
17	29 030 309 635 705 054	0	4 418 118 886 987 754 341 770
18	361 749 878 496 079 778	87 106 950 271 042 689 032	121 913 396 076 344 218 930 045
19	4 459 396 682 866 920 534	0	3 238 352 620 436 399 748 512 108
20	55 391 169 255 983 979 555	24 503 579 727 182 933 530 758	88 514 516 642 574 170 326 003 422

torized and parallelized. Because of this and the elimination of attrition, the method has a significant advantage over the traditional method of computer generation of self-avoiding walks which could be parallelized, but is difficult to vectorize.

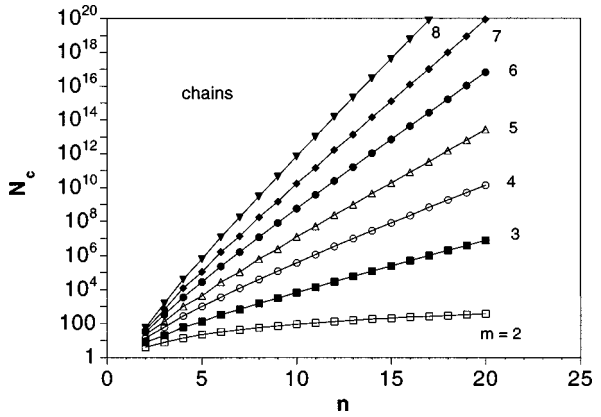

FIG. 10. The plot of the number of possible Hamiltonian chains N_c vs the length n for varying widths m of the $m \times n$ rectangle.

The main difficulty of the present method is its rapidly growing memory requirements. Consequently complete enumerations, especially for large numbers of points in 3D on the cubic lattice become difficult. We could, however, use the transfer matrix method for random sampling of the conformational space. We could randomly choose the connectivity states and the bond distributions and generate the compact conformations without attrition. Another major advantage of this method is that we may easily reduce the number of conformations by imposing constraints on the generated chains, for example from NMR NOEs we could know about various proximate pairs of residues. We might also fix one or both ends of the chain, or fix positions of some secondary structure elements of proteins, such as α helices or β sheets. Such imposed external constraints would significantly reduce the conformational space and enable the complete generation of conformations for much longer chains. The fixing of a certain point in a protein enables the immediate elimination of all conformations not compatible with the imposed constraint, thereby eliminating unneeded computations. The transfer method is powerful in this regard,

TABLE V. Enumerations of Hamiltonian walks (chains) on $m \times n$ rectangles.

n	$m=2$	$m=3$	$m=4$	$m=5$
2	4	8	14	22
3	8	20	62	132
4	14	62	276	1 006
5	22	132	1 006	4 324
6	32	336	3 610	26 996
7	44	688	12 010	109 722
8	58	1 578	38 984	602 804
9	74	3 190	122 188	2 434 670
10	92	6 902	375 122	12 287 118
11	112	13 878	1 128 446	49 852 352
12	134	29 038	3 342 794	237 425 498
13	158	58 238	9 767 588	969 300 694
14	184	119 518	28 217 820	4 434 629 912
15	212	239 390	80 709 424	18 203 944 458
16	242	485 822	228 864 620	80 978 858 522
17	274	972 414	664 060 262	333 840 165 288
18	308	1 960 830	1 800 346 140	1 456 084 764 388
19	344	3 923 326	5 002 457 832	6 021 921 661 718
20	382	7 882 494	13 825 549 136	25 904 211 802 080

n	$m=6$	$m=7$	$m=8$
2	32	44	58
3	336	688	1 578
4	3 610	12 010	38 984
5	26 996	109 722	602 804
6	229 348	1 620 034	12 071 462
7	1 620 034	13 535 280	175 905 310
8	12 071 462	175 905 310	3 023 313 284
9	82 550 864	1 449 655 468	43 551 685 370
10	572 479 244	17 198 428 572	682 958 971 778
11	3 808 019 582	142 545 533 336	9 735 477 214 522
12	25 304 433 030	1 580 868 297 042	144 397 808 917 246
13	164 452 629 818	13 246 916 541 978	2 033 155 413 979 838
14	1 062 773 834 046	139 620 415 865 920	29 105 375 742 858 518
15	6 777 328 517 896	1 183 338 916 049 852	404 654 754 079 984 324
16	42 944 798 886 570	11 997 107 474 280 224	5 656 098 437 704 094 140
17	269 706 791 277 978	102 719 325 162 193 010	77 710 312 229 803 403 554
18	1 683 956 271 732 804	1 010 824 101 911 587 178	1 067 886 114 091 399 967 842
19	10 445 800 698 724 066	8 728 784 450 632 453 306	14 517 649 840 508 475 301 004
20	64 470 330 298 173 718	83 947 749 266 911 632 982	196 974 144 293 101 997 656 968

since we may fix various structural elements at the same time.

We may easily generalize the transfer matrix method used for counting the compact Hamiltonian walks to the generation of compact conformations. For counting conformations it was useful to define connectivity states. This definition of ''states'' allows the unique generation of the chain only in the case of Hamiltonian circuits in 2D on the square lattice. In the case of circuits on the cubic lattice (as shown in a later paper) or Hamiltonian chains with ends a degeneracy problem is encountered. The sequence of connectivity states does not uniquely describe the conformation of a chain. The uniqueness of a conformation is specified only by alternating between the sequence of bond distribution states and connectivity states. Practically, in the generation of unique compact conformations it is useful to define in addition to the transfer matrix $\mathbf{T}$ describing the possibility of transitions between different connectivity states, another matrix $\mathbf{S}$ with elements S_{ij} showing the connectivity state obtained by the combination of the jth connectivity state and the ith bond distribution.

The transfer matrix method could also be easily generalized to irregular protein shapes. Equation (14) would then be replaced by

$$N_c = \mathbf{u}^T(\mathbf{T}_2\mathbf{T}_3\mathbf{T}_4 \ldots \mathbf{T}_{n-1})\mathbf{v}, \tag{18}$$

where the individual transfer matrices $\mathbf{T}_i$ for transitions between the ith and the $(i+1)$th columns differ in size but are specified to conform to the shape.

The new approach presented in this paper ought to become a standard useful method for studying globular proteins and compact polymers because of its large advantages over previous methods, including a recently published combinatorial algorithm based on two-matching and patching of bipartite graphs.[27]

One important application could be its use to obtain estimates of the chain entropies of proteins. Other direct appli-

cations to optimization problems such as the traveling salesman problem are also possible.

We have developed a new very efficient and attrition-free method for generating compact lattice conformations. This allows for the studies of the complete conformational space for larger lattice proteins, where traditional methods are restricted to sampling of the conformational space. The applicability of this new theoretical tool is currently under study.

ACKNOWLEDGMENTS

We are grateful to the National Research Council for the Senior Research Fellowship awarded to A. Kloczkowski.

[1] A. J. Guttmann and I. G. Enting, Phys. Rev. Lett. **76**, 344 (1996).
[2] N. Madras and G. Slade, *The Self-Avoiding Walk* (Birkhauser, Boston, 1993), pp. 393–394. The authors of this book combined values of the number of self-avoiding walks from many sources for self-avoiding walks on the square lattice up to 34 steps, and on the cubic lattice up to 21 steps. Longer walks on the square lattice, up to 39 steps were calculated by A. R. Conway, I. G. Enting, and A. J. Guttmann, J. Phys. A **26**, 1519 (1993). The newest enumerations of self-avoiding walks on the square lattice from 39 to 51 steps long are reported by those authors in A. R. Conway and A. J. Guttmann, Phys. Rev. Lett. **77**, 5284 (1996). Those terabyte-scale calculations required hours of mass parallel supercomputer time and 10 GB of memory.
[3] J. des Cloizeaux and G. Jannink, *Polymers in Solution* (Oxford, New York, 1989).
[4] A. D. Sokal, in *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, edited by K. Binder (Oxford, New York, 1995), pp. 47–124.
[5] P. J. Flory, *Statistical Mechanics of Chain Molecules* (Interscience, New York, 1969).
[6] H. S. Chan and K. A. Dill, J. Chem. Phys. **90**, 492 (1989).
[7] H. S. Chan and K. A. Dill, J. Chem. Phys. **92**, 3118 (1990).
[8] H. S. Chan and K. A. Dill, J. Chem. Phys. **95**, 3775 (1991).
[9] H. S. Chan and K. A. Dill, Macromolecules **22**, 4559 (1989).
[10] E. I. Shakhnovich and A. M. Gutin, J. Chem. Phys. **93**, 5967 (1990).
[11] E. I. Shakhnovich and A. M. Gutin, Nature (London) **346**, 773 (1990).
[12] E. I. Shakhnovich, G. Farztdinov, A. M. Gutin, and M. Karplus, Phys. Rev. Lett. **67**, 1665 (1991).
[13] D. A. Hinds and M. Levitt, Proc. Natl. Acad. Sci. USA **89**, 2536 (1992).
[14] D. G. Covell and R. L. Jernigan, Biochemistry **29**, 3287 (1990).
[15] R. L. Jernigan and D. G. Covell, in *Proteins: Structure, Dynamics and Design*, edited by V. Renugoplakrishnan (ESCOM, Leiden, 1991), pp. 346–351.
[16] R. L. Jernigan, Curr. Opin. Struct. Biol. **2**, 248 (1992).
[17] R. L. Jernigan and K.-L. Ting, in *Statistical Mechanics, Protein Structure, and Protein Substrate Interactions*, edited by S. Doniach (Plenum, New York, 1994), pp. 317–326.
[18] I. Bahar and R. L. Jernigan, Biophys. J. **66**, 454 (1994); **66**, 467 (1994).
[19] T. Ishinabe and Y. Chikahisa, J. Chem. Phys. **85**, 1009 (1986).
[20] V. S. Pande, C. Joerg, A. Y. Grosberg, and T. Tanaka, J. Phys. A **27**, 6231 (1994).
[21] A. Kloczkowski and R. L. Jernigan, Comput. Theor. Polymer Sci. **7**, 163 (1997).
[22] S. Hemmer and P. C. Hemmer, J. Chem. Phys. **81**, 584 (1984).
[23] B. Derrida, J. Phys. A **14**, L5 (1981).
[24] D. J. Klein, J. Stat. Phys. **23**, 561 (1980).
[25] T. G. Schmalz, G. E. Hite, and D. J. Klein, J. Phys. A **17**, 445 (1984).
[26] A. Kloczkowski and R. L. Jernigan, Macromolecules **30**, 6691 (1997).
[27] R. Ramakrishnan, J. F. Pekny, and J. M. Caruthers, J. Chem. Phys. **103**, 7592 (1995).